DOCUMENT RESUME

ED 363 662 TM 020 784

AUTHOR Bushery, John M.; And Others
TITLE The Schools and Staffing Survey: How Reinterview Measures Data Quality.
PUB DATE [93]
NOTE 7p.; Author is affiliated with the National Center for Education Statistics (NCES).
PUB TYPE Reports - Descriptive (141) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Administrator Characteristics; Comparative Testing; Elementary Secondary Education; *Followup Studies; *Interviews; *Mail Surveys; *National Surveys; Questionnaires; Reliability; *Research Methodology; Responses; School Personnel; *School Surveys; Secondary School Teachers; Teacher Characteristics; Teacher Supply and Demand; Telephone Surveys
IDENTIFIERS Data Quality; *Schools and Staffing Survey (NCES)

ABSTRACT

The first mail reinterview ever conducted by the Census Bureau was conducted as part of the Schools and Staffing Survey (SASS) of 1991. Because there were some questions about the reliability of items from the 1988 SASS reinterview, the 1991 study was examined by a mail reinterview. The methodology and responses are discussed, and response variance rates for questions reinterviewed in the 1988 and 1991 SASS cycles were compared. Mail reinterview respondents displayed lower simple response variance than under the telephone-telephone procedure. Findings suggest that mail reinterview respondents are cooperative and careful, although they may leave more difficult or uncertain questions blank. Implications of these findings are discussed. Mail surveys may provide data that are as good or better than some surveys now conducted by telephone or in person. Future survey plans are reviewed. Six tables present study findings. (Contains 7 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

TM

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED 363 662

THE SCHOOLS AND STAFFING SURVEY: HOW REINTERVIEW MEASURES DATA QUALITY

John M. Bushery and Daniel Royce, Bureau of the Census
and Daniel Kasprzyk, National Center for Education Statistics

KEY WORDS: Data quality, reinterview

1. INTRODUCTION

The Schools and Staffing Survey (SASS) is a good example of how a reinterview program can contribute to improved data quality by identifying questions which need improvement. We believe we have improved one aspect of SASS data quality, simple response variance -- in part because the SASS reinterview program identified questions needing improvement.

The 1991 SASS reinterview results also suggest that mail respondents provide more reliable data than those interviewed in a telephone follow-up operation.

1.1 The SASS Surveys

The National Center for Education Statistics (NCES) sponsors, and the U.S. Bureau of the Census conducts, the Schools and Staffing Survey (SASS) to provide data on teachers, school administrators, schools, and local education agencies.

The SASS runs on a three-year cycle, the first in 1987-88 and the second in 1990-91. The Census Bureau conducts the SASS by mail, with telephone follow-up of cases not responding by mail.

Mail response rates range from 49 percent (for private schools) to 80 percent (for public school administrators), with final response rates between 83 (private school teachers) and 97 percent (public school administrators again). We completed one-sixth to one-third of the cases using telephone follow-up.

1.2 The SASS Reinterview Program

Two major purposes of reinterview programs are quality assurance and estimating response error [1].

The SASS reinterviews estimate simple response variance, a measure of the inconsistency between responses over repeated applications of a question. Our main goal is to identify questions needing improvement for the next cycle of SASS. We identify problematic questions in the reinterview and follow up with cognitive research and other questionnaire design techniques to make the improvements.

To estimate response variance accurately, the survey error model assumptions require the reinterview to be an independent replication of the original interview. Independence is difficult to achieve because the respondent might remember his or her answer to the original interview question. To the extent a reinterview lacks independence, response variance may be underestimated. Operational constraints often make it difficult or impossible to conduct the reinterview as an exact replication of the original interview. When a reinterview does not replicate the original interview perfectly, the differences in methodology may overstate the response variance.

The SASS reinterviews fail to replicate the original interview in two respects:

- All SASS reinterviews contained fewer questions than their original counterparts.
- The original SASS surveys used self-administered mail-return questionnaires (with telephone follow-up of non-respondents). Except for the 1991 SASS School Survey, all the reinterviews were conducted by telephone.

We conducted the Census Bureau's first-ever mail reinterview in the 1991 SASS School Survey. Some of the 1988 SASS reinterview findings suggested that for some questions, the reinterview model assumptions were not adequately met [2]. Section 2.3 discusses this topic in more detail. These results prompted us to evaluate the 1991 SASS School questionnaire through a mail reinterview.

1.3 Response Variance Measures

Response error consists of response variance and bias. The Census Bureau estimates two main metrics (from unweighted data) to quantify response variance, the **gross difference rate** and the **index of inconsistency**. In a categorical variable, one-half the gross difference rate equals the simple response variance. The gross difference rate also represents the proportion of respondents who change their answers from one interview to the next. In a question with a gross difference rate of 20 percent, one fifth of the respondents changed their answers.

The index of inconsistency is a relative measure of response variance. A simplified definition of the

TM020784

index is the ratio of the simple response variance to the total variance of a characteristic. The L-fold index of inconsistency is a weighted average of the indices over all categories in a multi-category question. An index of 50 means that half the total variance of a characteristic can be attributed to response variance. Experience provides a rough rule of thumb for interpreting the index of inconsistency. If the index is:

- less than 20, response variance is low.
- between 20 and 50, response variance is moderate.
- greater than 50, response variance is high.

High response variance means the question itself causes at least as much of the variability in the data as the variability among respondents in the population. Two reasons for high response variance are:

- The question is poorly worded and confuses the respondent.
- The information requested is too difficult for the respondent to provide.

Because the index of inconsistency estimates the ratio of two variances, the index itself has high variability. If the data don't provide enough cases in each original-by-reinterview outcome cell, a reliable estimate of the index cannot be computed.

2. REINTERVIEW RESULTS

This paper compares response variance results for questions reinterviewed in both the 1988 and 1991 cycles of SASS. Table 1 shows reinterview sample sizes and completion rates for 1988 and 1991. We used unweighted data and tested all comparisons at $\alpha = 0.10$. Tables 3 through 6 display 90 percent confidence intervals in parentheses.

The Administrator and Teacher surveys ask both attitudinal and factual questions. In 1988 the attitudinal questions we reinterviewed showed high levels of inconsistency [2]. Inconsistency in attitudinal questions may result from simple response variance or from actual changes in attitudes between the original interview and reinterview. In 1991, we decided to concentrate the reinterview on factual questions -- with the aim of improving future cycles of the SASS.

In the 1988 SASS, we could estimate the index of inconsistency reliably for 35 of the 45 factual questions we reinterviewed. We estimated the index reliably for 109 of the 126 factual questions reinterviewed in 1991 [3]. Table 2 summarizes the results of both SASS reinterviews.

Keep in mind that the distributions in Table 2 are not strictly comparable. We purposively selected different sets of questions for the two reinterviews. We evaluated 15 factual questions common to both cycles of SASS. Eleven of these questions received significant revisions in 1991. Four of the revised questions displayed reduced response variance. Our question improvement efforts have paid off, at least partially.

Table 1. SASS Reinterview Sample Sizes

	1988	1991
Administrator Survey		
Eligible for Reinterview	1309	1123
Response Rate	87%	87%
Teacher Survey		
Eligible for Reinterview	1126	1101
Response Rate	75%	74%
School Survey		
Eligible for Reinterview	1309	1123
Overall Response Rate	87%	84%
Attempted Mail Reinterview		50%
Percent Completed by Mail		43%
Attempted Telephone Reinterview *		57%
Percent Completed by Telephone		41%

* Includes 80 reinterviews not returned by mail and 85 original mail interviews returned too late for mail reinterview.

Table 2. Summary of SASS Reinterview Results *

Response Variance	1988	1991
All Three Components		
Low	4 (11%)	43 (39%)
Moderate	14 (40%)	38 (35%)
High	17 (49%)	28 (26%)
Administrator and Teacher Surveys		
Low	4 (19%)	26 (36%)
Moderate	8 (38%)	26 (36%)
High	9 (43%)	21 (29%)
School Survey		
Low	0 (0%)	17 (47%)
Moderate	6 (43%)	12 (33%)
High	8 (57%)	7 (19%)

* Questions for which index could be reliably estimated.

2.1 Administrator and Teacher Survey Results

The two Administrator questions reinterviewed in both SASS cycles ask whether the respondent earned a bachelor's degree and a master's degree. These "degree earned" questions are virtually the same as the corresponding Teacher survey questions. The results for Administrators were nearly identical to the

Teacher results shown in table 3. The 1988 question provided a list of possible degrees and asked the respondent to "mark all that apply." The 1991 question asked, "Do you have a bachelor's degree?" If "Yes," the next question asked "Do you have a master's degree?" The remaining degrees (associate, doctor's, etc.) used a "mark all that apply" approach. Table 3 suggests the direct question format produces more reliable data for degree earned.

Table 3. Teacher Survey Reinterview Results
-- Degrees Earned --

	1988	1991
Bachelor's Degree		
Percent Yes[1]	97.6	98.1
GDR *	7.5	0.6
	(6.0 - 9.2)	(0.3 - 1.3)
Index	79.5	Too few cases
	(64.2 - 98.5)	did not mention
Master's Degree		
Percent Yes[1]	41.5	41.4
GDR *	4.3	1.1
	(3.2 - 5.7)	(0.6 - 1.9)
Index *	8.9	2.2
	(6.7 - 11.8)	(1.2 - 3.9)
Professional Diploma / Ed. Specialist		
Percent Yes[1]	4.4	4.7
GDR	7.0	5.2
	(5.6 - 8.7)	(4.1 - 6.8)
Index	69.8	62.7
	(56.0 - 87.1)	(48.2 - 81.6)
Associate Degree		
Percent Yes[1]	13.7	6.7
GDR	8.1	6.9
	(6.6 - 9.9)	(5.5 - 8.6)
Index	36.9	54.2
	(30.1 - 45.3)	(43.0 - 68.2)

[1] Responded "Yes" in original interview.
* Statistically significant difference between 1988 and 1991.

In the Teacher survey in both SASS cycles we also reinterviewed questions on teaching assignment, years in teaching, and plans to remain in teaching (an attitude type question). None of these questions exhibited significantly improved response variance.

The teaching assignment questions reinterviewed in 1988 and 1991 were similar but not strictly comparable. In 1991 we reinterviewed a screener question used to identify teachers, which asked about full and part-time status and included categories for itinerant teachers, long-term substitutes, other professional staff, and administrators (the last two are out of scope for the Teacher survey). The 1988 question simply asked about full-time and four levels of part-time teaching. The 1988 question includes all full-time teachers, the 1991 figure includes only regular full-time teachers. These design differences make it difficult to compare the two questions, but response variance on the number of full-time teachers showed no significant change between 1988 and 1991. The new categories seem to cause respondents some uncertainty -- about six percent (s.e. = 0.8) of the respondents described their assignment as itinerant teacher, long-term substitute, other professional staff, or administrator in the original interview. Only three percent (s.e. = 0.6) selected one of these answers in the reinterview. The data suggest the "itinerant teacher" category is the main source of this inconsistency. It may help to define "itinerant" more clearly on future questionnaires.

The 1988 "years teaching" questions asked, "... how many years have you worked as a full-time teacher in public and/or private schools ..." (repeated for part-time) and provided a cross-tabulation for the respondent to complete:

	Years full-time	Years part-time
Public		
Private		

In 1991 we changed the format to ask four separate questions:

- " ... how many years have you worked as a full-time teacher in private ...,"
- "... part-time in private ...,"
- "... full-time in public ...," and
- "... part-time in public ..."

Table 4. Teacher Survey Reinterview Results
-- Years Teaching --

	1988	1991
Full-time, Public		
GDR	7.6	7.0
	(6.1 - 9.5)	(5.5 - 8.9)
L-fold Index	10.8	9.8
	(8.7 - 13.4)	(7.7 - 12.4)
Part-time, Public		
GDR	9.0	6.6
	(6.7 - 12.0)	(5.0 - 8.6)
L-fold Index	44.4	42.5
	(33.2 - 59.3)	(32.5 - 55.7)
Full-time, Private		
GDR	5.2	5.3
	(3.6 - 7.4)	(3.3 - 8.7)
L-fold Index	12.4	8.8
	(8.7 - 17.7)	(5.4 - 14.4)
Part-time, Private		
GDR *	3.4	7.5
	(2.1 - 5.8)	(4.8 - 11.6)
Index	38.5	37.8
	(23.0 - 64.4)	(24.4 - 58.4)

* Statistically significant difference between 1988 and 1991.

BEST COPY AVAILABLE

We grouped the responses into the four categories of interest to the NCES:

- less than three years,
- three to nine years,
- 10 to 20 years,
- more than 20 years.

Unfortunately, no improvement resulted. The full-time estimates enjoyed low response variance in both years, and the part-time estimates exhibited moderate response variance in both cycles of SASS (Table 4.).

The final Teacher question reinterviewed in both SASS cycles was, "How long do you plan to remain in teaching?" The consistency of this attitude-type question decreased between 1988 and 1991. The gross difference rate increased from 39.5 percent (36.8% - 42.6%) to 46.8 percent (44.0% - 49.9%) and the L-fold index increased from 55.4 (51.6 - 59.6) to 66.6 (62.6 - 71.1). Since we did not change this question, we speculate that teachers' attitudes in 1991 were less stable than in 1988.

Increased response variance among public school teachers drove the overall decrease in consistency -- private school teachers showed no significant change in response variance between 1988 and 1991

2.2 School Survey Results

In the School survey, we reinterviewed four questions in both 1988 and 1991. Although these questions were virtually unchanged between the two cycles, they showed a small but statistically significant decrease in response variance.

We think a better replication of the original interview by the reinterview in 1991 caused some of this decrease. Table 5 shows the reinterview results for these questions.

The question, "Which best describes the community in which this school is located?" contained ten categories in 1988 and 1991.

1 rural or farming community
2 small city or town, not a suburb of a large city
3 medium-sized city
4 suburb of medium city
5 large city
6 suburb of large city
7 very large city
8 suburb of very large city
9 military base or station
10 Indian reservation

The index of inconsistency for these categories ranged from 21.1 to 68.8 in 1988 and from 22.2 to 62.1 in 1991. The overall response variance (L-fold index) for this question improved slightly, but remains in the moderate range. "Community" is an important variable in the NCES' analyses. Fortunately, the NCES is now able to obtain this information from geographic data files [6], instead of asking the schools. The result will be more accurate data, with reduced respondent burden.

We reinterviewed three questions about programs offered by the school, "Which of the following programs and services are available to students in this school, either during or outside of regular school hours, and regardless of funding source -

- bilingual education
- English as a second language
- extended day or before-or-after-school day-care."

Table 5. School Survey Reinterview Results

	1988	1991
Which best describes the community in which this school is located?		
GDR *	34.7	30.4
	(32.3 - 37.1)	(27.9 - 32.9)
L-fold Index *	42.4	37.6
	(39.6 - 45.4)	(34.7 - 40.9)
bilingual education		
Percent Yes[1]	15.3	14.2
GDR *	16.2	12.1
	(14.5 - 18.2)	(10.5 - 14.1)
Index	53.5	45.1
	(47.7 - 60.0)	(39.0 - 52.3)
English as a second language		
Percent Yes[1]	31.6	28.3
GDR *	16.1	13.7
	(14.4 - 18.1)	(12.0 - 15.8)
Index *	37.1	30.1
	(33.1 - 41.7)	(26.3 - 34.6)
extended day or before-or-after-school day-care		
Percent Yes[1]	16.3	23.0
GDR	9.3	8.8
	(7.9 - 11.0)	(7.4 - 10.6)
Index	31.7	24.7
	(26.8 - 37.4)	(20.5 - 29.7)

[1] Responded "Yes" in original interview.
* Statistically significant difference between 1988 and 1991.

2.3 Mail versus Telephone Results (1991)

In 1991 we revised the School survey reinterview procedures:

- We used a mail reinterview for mail respondents and a telephone reinterview for telephone follow-up cases.
- We requested the same respondent complete the

BEST COPY AVAILABLE

reinterview questions as answered the original School survey.

Both procedural changes helped the reinterview replicate the original survey better. We decided to specify the original school respondent as the reinterview respondent, because in the 1988 reinterview we inadvertently changed the reinterview's respondent-selection rules by combining the Administrator and School reinterview questionnaires. We suspect many administrators had an assistant or secretary complete the original School survey. Changing respondents between the original and reinterview tends to overstate response variance in the 1988 School survey.

We did not conduct a controlled experiment, but reinterviewed by mail whenever possible and by telephone when necessary, obtaining about 465 mail-mail cases and 270 telephone-telephone cases. This analysis covers the same four School survey questions discussed in section 2.2. Under the mail-mail procedure almost all the School questions reinterviewed in 1991, including the four in Table 6, displayed lower simple response variance than under the telephone-telephone procedure.

Table 6. Mail Original/Reinterview versus Telephone Original/Reinterview

	Mail-Mail	Telephone-Telephone
Community School Located		
GDR *	19.0	39.9
	(16.3 - 22.2)	(35.5 - 45.2)
L-fold Index *	24.0	48.6
	(20.6 - 28.2)	(43.2 - 55.1)
Bilingual Education		
GDR *	6.9	18.6
	(5.2 - 9.1)	(15.2 - 23.0)
Index *	31.5	55.3
	(23.5 - 42.0)	(45.3 - 68.2)
English as 2nd Language		
GDR *	10.9	15.7
	(8.8 - 13.6)	(12.6 - 19.8)
Index	24.2	33.5
	(19.6 - 30.1)	(26.8 - 42.3)
Extended Day Care		
GDR *	6.7	11.5
	(5.1 - 8.9)	(8.8 - 15.2)
Index *	19.7	31.9
	(14.7 - 26.4)	(24.5 - 42.2)

* Statistically significant difference between mail-mail and telephone-telephone.

We observed lower response variance in both numerical data (for example, head counts of students enrolled) and non-numerical data. Royce [3] details results for all School survey questions reinterviewed in 1991. We can think of four possible reasons for this result.

- Only respondents who answered the original survey by mail were eligible for the mail reinterview. These respondents were likely to be more cooperative and answer the questions more carefully in both interviews.
- Respondents interviewed by mail may take time to look up the answers to questions from records or they may go through a more careful, but more lengthy, thought process to provide the needed facts. In contrast, those interviewed by telephone may feel the interviewer prefers a speedy response to an accurate one, so give their "best guess-timate." Research has shown some respondents employ what survey practitioners call "satisficing." [4] In satisficing, the respondent expends just enough effort to satisfy the interviewer. Also, respondents interviewed by telephone may not feel free to take the time to look up records while the interviewer is waiting on the phone [5].
- Mail respondents may leave more difficult or uncertain questions blank. The Census Bureau's interviewers work very hard to get responses to all questions. An interviewer may manage to obtain an answer to a difficult question, but an unreliable answer. Mail respondents, on the other hand, may simply leave that question blank. We have found higher item non-response among the mail returns than in the telephone follow-up cases.
- Mail respondents may photocopy the original questionnaire after completing it and refer to their original answers when completing the mail reinterview.

We think some combination of the first three explanations is the most reasonable. Mail respondents, by definition, are more cooperative and motivated than those we must follow-up by telephone. And mail interviewing probably promotes more careful responses and more use of records.

We eliminated the last possibility. Mail respondents using photocopies of their original interviews can account for only a small part of the mail-mail versus telephone-telephone differences. We concluded that only a small fraction of the mail reinterview respondents might have used photocopies, and that these cases had little effect on the response variance differences between the two procedures. We hypothesized that respondents using photocopies would give consistent answers to all questions in the reinterview. We discarded all cases where the first 11 of the 21 reinterview questions matched. These cases accounted for only 6.5 percent of the reinterview sample and had only a negligible effect on the comparisons.

These findings on the quality of mail response data have implications beyond the SASS. Perhaps mail surveys can provide as good or better data than some surveys now conducted by telephone or in person -- and at lower cost. For the SASS, we need to determine whether the more consistent data achieved through mail results from the type of respondent who answers by mail and whether increased item non-response will cancel the gains of improved consistency.

3. PLANS FOR THE FUTURE

Reinterview programs can be a valuable diagnostic tool to identify questions which need improvement, or which perhaps should be dropped. The NCES and the Census Bureau are committed to producing accurate and reliable SASS data. They have heeded the reinterview's diagnosis and have acted to make improvements -- with some success.

What about the future? Both agencies are firmly committed to developing a first-class survey. The 1992 Teacher Follow-up Survey (TFS), which surveyed a subsample of 1991 SASS teachers, used a probing, reconciled reinterview to learn the reasons for inconsistent responses. We hope not only to identify the less reliable questions, but to gather information about why inconsistencies occur.

Plans for the future include:

- Focus at least some cognitive research on the reinterview findings.
- Consider using reconciled, probing reinterviews in the SASS to learn more about why inconsistencies occur.
- Consider expanding the mail reinterview to the Teacher and Administrator surveys.
- Apply quality assurance methods to data collection.
- Reinterview small, non-random samples to solve specific data quality problems, for example unacceptably high pre-edit rejects.
- Use reinterview methods to evaluate coverage in teacher listings (the frame of the SASS teacher sample).
- Maintain a strong commitment to a continual cycle of evaluation and improvement of SASS questionnaires, methods, and procedures.

REFERENCES

[1] Forsman, G. and Schreiner, I., "The Design and Analysis of Reinterview: An Overview," in Biemer, P. P., et al., editors, Measurement Errors in Surveys, 1991, John Wiley & Sons, Inc., pp. 280-281.

[2] Newbrough, J., "Report of the SASS-2/3(R) and SASS-4(R) Reinterview," June 28, 1989, internal Census Bureau Report.

[3] Royce, D., "1991 Schools and Staffing Survey (SASS) Reinterview Response Variance Report," July 1992, internal Census Bureau Report.

[4] Krosnick, J. A., "The Impact of Satisficing on Survey Data Quality," Proceedings of the 1990 Annual Research Conference, U.S. Bureau of the Census, March, 1990, pp. 835-845.

[5] Dillman, D. and Tarnai, J., "Mode Effects of Cognitively Designed Recall Questions: A Comparison of Answers to Telephone and Mail Surveys," in Biemer, P. P., et al., editors, Measurement Errors in Surveys, 1991, John Wiley & Sons, Inc., p. 76.

[6] Johnson, F., "Assigning Type of Locale Codes to the 1987-88 CCD Public School Universe," National Center for Education Statistics Technical Report, July 1989.

[7] Parmer, R.J., Shen, P., and Tan, A.L., "Mail Versus Telephone Response in the 1991 Schools and Staffing Surveys," Joint Statistical Meetings, Boston, 1992.

This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributable to the authors and do not necessarily reflect those of the Census Bureau.